The global transcriptional regulatory network for metabolism in *Escherichia coli* exhibits few dominant functional states

Christian L. Barrett[1], Christopher D. Herring[2], Jennifer L. Reed[3], and Bernhard O. Palsson[1]

***Short Abstract* — A principal aim of systems biology is to develop mechanistically detailed *in silico* reconstructions of cells and to use them as dynamic models for the purpose of discovering and understanding cellular capabilities and phenotypes. We used the genome-scale reconstruction of the integrated metabolic and transcriptional regulatory network for *Escherichia coli*—composed of 1,010 gene products—to assess the properties of all functional states computed in 15,580 different growth environments. We observed that the integrated network is organized to achieve few dominant cellular phenotypes. This result demonstrates how a bottom-up approach can elucidate an organism's preferred environments and functional capabilities in a mechanistically transparent way.**

***Keywords* — *in silico* integrated reconstructions, cellular phenotypes, dynamic simulations, mechanistic basis for genotype-phenotype relationship.**

I. Purpose

The long term goal of accurate *in silico* modeling of the dynamic behavior of cells is consistent with the systems approach to biological discovery and understanding. Progress towards this goal is demonstrated when computational models are used to integrate large amounts of molecular data to interpret and predict the range of cellular phenotypes. We investigated [1] the utility of the systems approach by using a genome-scale, literature-curated reconstruction of the integrated transcriptional regulatory and metabolic network for *Escherichia coli* (iMC1010v1) to computationally assess growth phenotypes in an exhaustive library of minimal-medium growth environments. iMC1010v1 accounts for a total of 1,010 ORFs in *E. coli*, representing 906 metabolic ORFs that enable 932 unique biochemical reactions (including transport reactions) among 625 metabolites and 104 transcription factors regulating the expression of 479 of the 1,010 ORFs in the reconstruction.

II. Results

We globally characterized the full range of functional states that the integrated transcriptional regulatory and metabolic network could exhibit by simulating network growth states in an exhaustive library of 15,580 minimal-media conditions. Each functional state was described by an activity profile that contained both the calculated expression state(s) of each gene and the logical interactions between all TFs and the genes they are known to regulate. After clustering the activity profiles and applying a standard dimensionality-reduction technique, we showed that (i) the set of all possible network states has few dominant modes; (ii) these modes are primarily organized according to the terminal electron acceptor and whether or not glucose or gluconate was the carbon source; (iii) the clusters can be differentiated on the basis of the activity of relatively few TFs, consistent with recent genome-scale experimental results and limited expression profiling data; and (iv) the integrated network gives a crisper clustering structure of functional states than without regulatory information.

III. Conclusion

Our study, on the whole, represents an integration of diverse sources of data. We found that the most detailed and informative state description used in our study led not to a more complicated view of the cell's "state space", but to a markedly simpler one. We interpret this result to be a reflection of the internal organization of the cell, which is that of a system robustly structured against all but a few characteristic behaviors. This result is consistent with theoretical work on "complexity for simplicity" and on regulatory network state attractors. We expect that this reflection will become increasingly pronounced as larger amounts of additional types of state information are further integrated into a mechanistic format.

References

[1] Barrett CL, Herring C, Reed JL, Palsson BO (2005) The global transcriptional regulatory network for *Escherichia coli* exhibits few dominant functional state. *PNAS* **102(52)**, 19103-19108.

[1]Department of Bioengineering, University of California, San Diego. E-mail: {cbarrett, bpalsson}@bioeng.ucsd.edu

[2]Mascoma Corpoaration. E-mail: cherring@mascoma.com

[3]Department of Chemical and Biological Engineering, University of Wisconsin, Madison. E-mail: reed@engr.wisc.edu